

Abstract

The invention applies a probabilistic approach to combining evidence regarding the correct classification of items. Training data and machine learning techniques are used to construct probabilistic dependency models that effectively utilize evidence. The evidence includes the outputs of one or more classifiers and optionally one or more reliability indicators. The reliability indicators are, in a broad sense, attributes of the items being classified. These attributes can include characteristics of an item, source of an item, and meta-level outputs of classifiers applied to the item. The resulting models include meta-classifiers, which combine evidence from two or more classifiers, and tuned classifiers, which use reliability indicators to inform the interpretation of classical classifier outputs. The invention also provides systems and methods for identifying new reliability indicators.